On reliability and trustworthiness of high-level fusion-based decision support systems: basic concepts and possible formal methodologies

Per Svensson
Swedish Defence Research Agency, Stockholm, Sweden
Per.Svensson@foi.se

Abstract - The paper summarizes the result of a literature study of robust uncertainty management methodologies, carried out to indicate options available in the design and construction of trustworthy decision support systems based on high-level information fusion methods. Among the candidate methodologies briefly discussed for creating trustworthy decision support are robust Bayesian statistics, imprecise probabilities and sensitivity analysis of simulation models. However, few reports of the application of such techniques in information fusion software systems were found.

Keywords: robust Bayesian statistics, imprecise probabilities, sensitivity analysis, decision support systems, information fusion

1 Introduction

An important topic in information fusion research is fusion performance, including *relevance, trustworthiness, reliability,* and *robustness.* Llinas [1], *e.g.*, lists relevance of decisions or recommendations, correctness in reasoning (reliability), and adaptability in reasoning (robustness) among representative assessment criteria for high-level fusion systems (i.e., situation and impact refinement, or “level 2 and 3”, systems), *cf.* [2, 3]. Sheridan [4] suggests reliability, robustness, familiarity, understandability, explication of intention, usefulness, and dependence as potential components in a user’s relation of trust in a command and control system.

In this paper we use the term *trustworthiness* as a human factors attribute of a system, meaningful only when given a specific usage domain and an adequate relationship between user, system, and environment, including sufficient user training and experience for the task at hand given the tool and an environment which stays within the specified usage domain of the system. The term *reliability* will be used to denote the technical system property of delivering quantitative results which are reasonably close to best possible, subject to known statistical error distributions. Thus, reliability is viewed as a technical precondition of trustworthiness. The term *robustness*, finally, is used to denote the property of a system to react appropriately to exceptional conditions, including to avoid making large changes in recommendations as a consequence of small changes in input data. The latter property will also be referred to as *stability.* In essence, the environment model of a robust DSS needs to be able to discover situations of high risk and quantify the range of recommendation uncertainty inherent in these.

While numerous papers deal with trust and trustworthiness as human factors issues in command and control and intelligence management systems, *e.g.* [5, 6], few studies discuss basic reliability concepts in high-level fusion-based decision support systems (HLFB-DSS) and applications. In such systems, generic conceptual architectures such as Bayesian networks [7], decision trees, influence diagrams [8], and Bayesian games [9] are frequently adopted, while different approaches to uncertainty management are often combined [2, 3, 10]. However, the effects of such combination on precision and robustness of the output are rarely fully understood. Consequently, current approaches to information fusion methodology are not likely to be sufficient to satisfy HLFB-DSS users’ reliability requirements. Indeed, unless concepts and methodologies are found and generally applied which enable researchers and developers to achieve and demonstrate reliability of high-level information fusion methods and algorithms, operational decision makers are unlikely to be willing to trust or use decision support systems based on such techniques.

On the other hand, in recent decades substantial research and engineering work in robust decision analysis and its applications has been performed in other fields. Since this partly fundamental research indicates that reliable behaviour of HLFB-DSS might be achievable, we believe that it needs to be noticed, evaluated and, if possible, adapted for use in the design of HLFB-DSS. With the aim of contributing to the initiation of such efforts, this paper discusses aspects of reliability and trustworthiness of certain classes of methods and models of decision-making under uncertainty. Such aspects are likely to be limiting factors in many complex decision-making and risk management situations. However, although we strongly believe that reliability assessment of HLFB-DSS needs to be based on well-understood theories for management of uncertainty, for fundamental reasons guaranteed performance is likely to remain an unachievable ideal, except perhaps in the simplest cases.

The paper is structured as follows. In Section 2, we introduce concepts and some applications of decision support methodology in a Bayesian framework. Next, we review three methodological approaches to increased reliability and robustness in DSS: *robust Bayesian analysis* (Section 3), *sensitivity analysis* (Section 4), and *imprecise probabilities*, including interval probabilities (Section 5). In Section 6, we review a paper by Arnborg [11], further elaborated in [12], on the relationship between robust Bayesian analysis and evidence theory, which concludes that these approaches may lead to quite different conclusions, at least in input-sensitive, *ill-conditioned* [13] [14] cases. In Section 7, we review experimental results presented in [15] on the application of various imprecise probability methods to the so-called Sandia challenge problems, a set of simple, abstracted models of expert judgment aggregation problems. Section 8 concludes the paper.

2. Decision analysis: concepts and applications

French and Rios-Insua [16] summarize the Bayesian approach to inference and decision analysis as a multistep process:

1. Modeling beliefs about a parameter of interest through a *prior* which, in presence of further information, is updated to the *posterior.*
2. Modeling *preferences* and *risk attitudes* about (possibly multicriteria) consequences with a (multiattribute) function.
3. Associate with each alternative its (multiattribute) *posterior expected utility.*
4. Propose the alternative which maximises the posterior expected utility.

Since the assessment of beliefs and preferences is in general a difficult task, and since the decision models needed in practice may be complex, the need for evaluating the sensitivity of the output with respect to changes to the inputs (*model*, *beliefs* and *preferences*) is typically great. In addition, since in the Bayesian framework the decision maker's judgments are encoded into inputs to the analysis, he or she is likely to want to explore their implications, as well as repeat the analysis process using revised judgments, until the space of alternative outcomes has been sufficiently explored and understood. Only then should a decision be made. Classical texts describing Bayesian decision analysis are [17][18].

In an influential paper, Apostolakis [19] discussed the concept of Bayesian analysis in probabilistic safety assessment (PSA) of technological systems. The purpose of doing a PSA is to make decisions regarding the safe operation of a facility. Expected utility theory provides the framework within which decisions can be analyzed in a formal manner and in accordance with several reasonable principles [18]. Because the events or phenomena of interest in PSA are usually very rare, thus lacking significant statistical or experimental support, the opinions of experts, and how to *elicit* prior probability estimates from them, acquire great significance.

In HLFB-DSS, a role somewhat analogous to that played by experts in PSA applications is filled by sensor subsystems which need to be subjected to extensive off-line experimentation and measurement, as well as modeling and model training (see, *e.g.*, [20, 21, 22]), before they can provide reliable, calibrated inputs to automatic fusion algorithms. In these calibration processes, elicitation of expert knowledge often plays a key role. High-level information fusion systems also frequently involve automatic interpretation of observations according to some more or less predefined behavioral model, *e.g.*, a model of the opponent's tactical doctrine, if one exists and is known. Ultimately, users may want to engage in a *mixed-initiative* dialog [23] with the DSS by providing prior probability estimates of parameters in such models.

Jiménez *et al.* [24] describe a generic PC-based decision support system based on a Bayesian multiattribute utility model that is intended to ease many of the operational difficulties involved in the decision analysis cycle. The system, GMAA, accounts for uncertainty about the alternative consequences, which can be defined in terms of continuous uniformly distributed intervals for each attribute, and provides several types of sensitivity analysis. It admits incomplete information about preferences by permitting interval responses to the probability questions put to the user, which leads to sets of utility functions and weight intervals. This makes the system suitable also for group decision support, where individual conflicting views in a group of decision makers can be captured through imprecise answers. The system was designed to be useful in a range of complex decision-making problems discussed in the operations research literature [25], such as military systems acquisition processes, analysis of alternatives for the disposition of surplus weapons-grade plutonium, *etc*.

Ekenberg and Thorbiörnson [26] provide a theory for analyzing decisions under risk when the available information is vague and imprecise, *cf.* Section 5. Partly based on this work, Larsson *et al.* [27] present a decision tree evaluation method for analyzing multiattribute decisions under risk. Information is modeled using convex sets of utility and probability measures restricted by closed intervals.

3 Robust Bayesian analysis

According to Berger [28], there is a common perception that foundational arguments lead to subjective Bayesian analysis as the only *coherent* method of behavior. This is indeed true if it is assumed that one can make arbitrarily fine discriminations in judgment about unknowns and utilities. In reality, however, it is very difficult to discriminate between, say, 0.10 and 0.15 as the subjective probability, *P(E)*, to assign to an event *E*. However, realistic foundational structures, or axiomatic theories, exist (see, *e.g.*, [29], and further references in [28]), based on axiomatics of behavior which acknowledge that arbitrarily fine discrimination is impossible. The conclusion of these theories is that coherent behavior corresponds to having *sets* of input models, priors, and utilities, which yield a range of possible Bayesian answers. If the range of answers is too large, the question may not be settled. Conceptually introduced in the 50´s by Good and later refined by Kadane and Chuang [30], modern surveys of this so-called *robust Bayesian analysis* are given in [28][31][32].

Berger *et al.* [32] declare the early 90's the golden age of robust Bayesian analysis. Since then, the need to consider Bayesian robustness has in fact increased dramatically, since models that are now routinely used in Bayesian analysis are sometimes so complex that their inputs (such as priors) can be elicited only in a casual fashion. Still, robust Bayesian methods have not yet had much impact on applications, although the already cited paper [24] by Jiménez *et al.* provides some examples. On the other hand, new research opportunities are offered by developments in algorithms, the possibility of using MCMC methods, and the growing need for robust behavior and sensitivity analysis in applications.

Robust Bayesian analysis provides tools to check the impact of the utility function, the prior and the model on the optimal decision alternative and its posterior expected utility. Three main approaches to Bayesian robustness can be distinguished: *informal*, *global sensitivity*, and *local sensitivity*. According to [32], a number of theoretical results show that one may model imprecision in beliefs and preferences through a *set of probability distributions* and a *set of utility functions*. These results have two basic implications. First, they provide a qualitative framework for sensitivity analysis in decision analytic problems, describing under what conditions one may undertake the standard approach of perturbing the initial probability-utility assessments. Second, they point to the basic concept of robust approaches: determining the set of *non-dominated alternatives*. In the common cases when the non-dominated set is too large to imply a final decision, one should look for additional information that would help reduce its size.

As indicated above, the usual practical motivation of robust Bayesian analysis is the difficulty in assessing the prior distribution. Instead, however, one could directly operate with the constraints $a_i \le Pr(s_i) \le b_i$ obtained by, *e.g.*, expert elicitation. If a parameterized utility function is assessed, the constraints are typically placed on the parameters of the utility, such as the risk aversion coefficient. In developing the model for the data itself there is a typically great imprecision, and a need for careful study of the model robustness. And when there are several decision makers and/or experts involved, it may not be theoretically possible to obtain a single model, prior, or utility; instead, one might be left with only sets of each, corresponding to differing expert opinions.

4 Sensitivity analysis: concepts and applications

Bayesian robustness is playing a role in SAMO (*Sensitivity Analysis of Model Output*), a network of researchers interested in investigating the relative importance of model input parameters on model predictions [34] in many applied areas. There are many reasons to check the sensitivity of the output (the optimal alternative) with respect to the inputs (model, beliefs and preferences). Since inputs to the analysis encode the decision makers' judgments, he or she should wish to explore their implications and possible inconsistencies. The need for sensitivity analysis is further emphasized by the fact that the assessment of beliefs and preferences is a difficult task.

Stability theory [35] provides a unifying sensitivity, or robustness, framework. When *strong stability* holds, careful elicitation should lead to decisions with expected utility close to the greatest achievable; when *weak stability* holds, at least one stabilized decision will have this property. However, when neither concept of stability applies, even small elicitation errors may lead to disastrous results in terms of large losses in expected utility.

Sensitivity analysis of simulation models [34][36][37] can be used to identify the most significant exposure or risk factors in a model, as an aid in identifying the important uncertainties for the purpose of prioritizing additional data collection or research, and it can play an important role in model verification and validation throughout the course of model development and refinement. Sensitivity analysis also can be used to provide insight into the robustness of model results when making decisions. In [37], a number of sensitivity analysis methods are surveyed.

Kleijnen [38] gives a survey on the use of statistical designs for "what-if" analysis in simulation, including sensitivity analysis, optimization, and validation/ verification. Sensitivity analysis is divided into two phases. The first phase is a pilot stage, which consists of screening or searching for important factors among possibly hundreds of potentially important factors. The second phase uses regression analysis to approximate the input/output transformation that is implied by the simulation model; the resulting regression model is also known as a *metamodel* or *response surface*. Regression analysis gives better results when the simulation experiment is well-designed, using either classical statistical designs (such as fractional factorials) or optimal designs. To optimize the simulated system, the analysts may apply *Response Surface Methodology* (RSM); RSM combines regression analysis, statistical designs, and steepest-ascent hill climbing. To validate a simulation model, again regression analysis and statistical designs may be applied.

In military applications of modeling and simulation, there are specified processes of verification, validation, and accreditation (VV&A) [38][39]. In a sense, information fusion may be seen as a special case of M&S, since information fusion is partly based on situation modeling using theories of uncertainty, and frequently

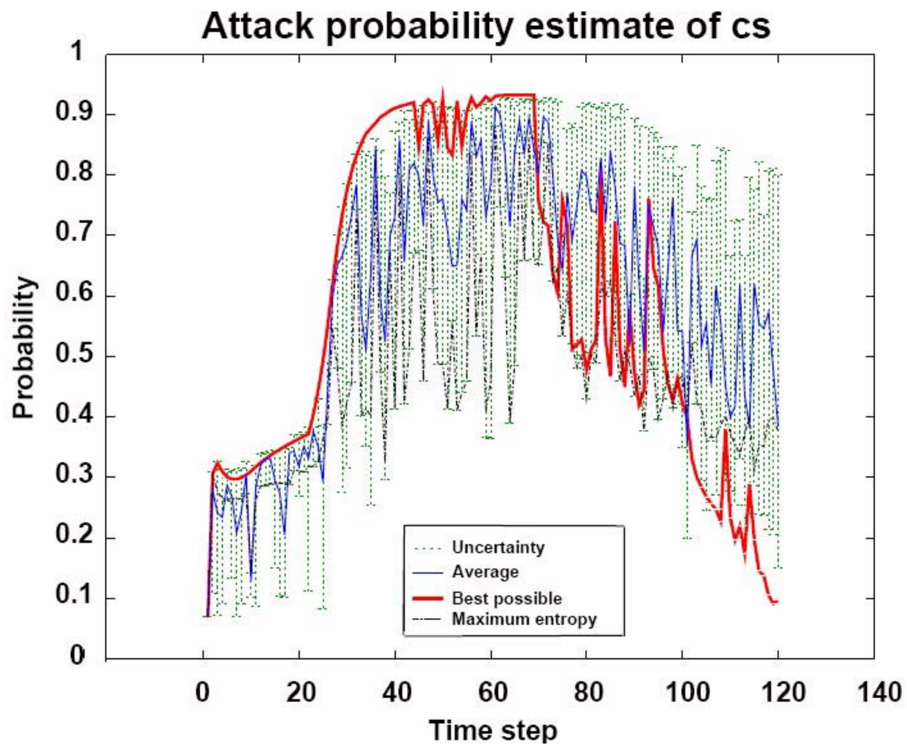

Figure 1. Result of a scenario-based simulation experiment for evaluating particle filter-based state estimation in a tactical plan recognition system using proactive control of sensor resources. From Johansson & Suzic [10].

employs Monte Carlo simulation techniques in model evaluation and decision-making. Including HLFB-DSS in the VV&A processes, a necessary provision for acceptance, therefore seems to be possible without having to introduce a great number of unfamiliar concepts.

In [10], and more thoroughly in [40], Johansson and Suzic study information acquisition for robust plan recognition, using Monte Carlo simulation in the form of particle filtering to obtain a measure of modeling uncertainty. According to [10], tactical commanders want to obtain predictive situation awareness. To do this effectively, they need real-time decision support tools which can both recognize basic tactical plans of the opponent, such as an imminent attack, and proactively control limited sensor resources by prioritizing dangerous plan alternatives in a sensible way. Johansson and Suzic introduce a particle filter that maintains a state estimate of opponent plans even when observations are lacking. The particle filter produces a multi-mode state representation with each particle as a mode.

The approach is studied in a scenario-based simulation experiment, see Figure 1. Initially, the suspected attacker (`cs`) is observed by both UAVs and observers on the ground. Estimated attack probability is close to best possible and the uncertainty is small. As estimated attack probability increases, the sensor control manages to keep the uncertainty relatively small by prioritizing this objective. Near the end of the scenario, attack probability has decreased (because `cs` is now moving away from the own targets) and the automatic sensor control gradually changes its priorities. Together with a built-in “gravity” bias of the model which pulls particles towards own targets, this explains why the uncertainty interval fails to cover the best possible estimate during the final time steps.

5. Imprecise probabilities

Recently, there has been considerable interest in theories of imprecise probabilities. The biannual conference series ISIPTA (*International Symposia on Interval Probabilities and Their Applications*) started in 1999 [41][42]. Another effort, of particular relevance for engineering applications, is the *Sandia Workshop on Alternative Representations of Epistemic Uncertainty*, held in August 2002. This workshop has been documented in a special issue of the journal *Reliability Engineering & System Safety* [43].

Imprecise probability is a generic term used to describe mathematical models that measure uncertainty without precise probabilities. This is certainly the case with robust Bayesian analysis, but there are many other imprecise probability theories, including in decreasing order of generality *coherent lower and upper previsions*, *coherent lower and upper, or interval, probabilities*, *Choquet capacities of order 2* (*cf.* Sec. 6 where this concept is defined and used), *belief and plausibility functions*, *possibility and necessity measures*, and *fuzzy logic*, see [29][44][45]. Thus, *e.g.*, belief functions are a special case of Choquet capacities of order 2 [29][11].

Some of these theories, such as fuzzy logic and belief functions, are only tangentially related to robust Bayesian analysis, while others are closely related. Seen from the robust Bayesian perspective of [32], the major difference between robust Bayesian analysis and these alternative theories is that robust Bayesian analysis stays with ordinary Bayesian intuition, considering bounded sets of individual prior distributions that are each compatible with prior beliefs. In contrast, the alternative theories view the sets themselves (not the individual priors) as the basic elements of the theory.

In [46], Walley claims that a general theory of imprecise probability can accommodate all the kinds of uncertainty and partial ignorance that are currently being studied, including vague or qualitative judgments of uncertainty, models for complete ignorance or near ignorance, random sets and multivalued mappings, and partial information about an unknown probability distribution. In [29] Walley describes the connection between robust Bayesian analysis and the theory of coherent lower and upper previsions (for a tutorial introduction to these concepts, see [47]).

According to Walley [46], coherent lower and upper previsions are needed in a general theory because these are direct generalizations of the most commonly used models (interval probabilities, Choquet capacities of order 2, belief functions, possibility measures, and probability distributions), so that a general theory of imprecise probability can be applied directly to these special models. Sets of probability distributions are also needed in a general theory because, at present, most examples of coherent models are presented in this form. This is the approach in the robust Bayesian theory, which uses a set of probability distributions as the canonical model for uncertainty. Since upper and lower envelopes of a set of probability distributions are always coherent upper and lower previsions, specifying a set of probability distributions is a simple way of constructing a coherent model. For example, after receiving new information, a set can be updated by using Bayes' rule to update each probability distribution in the set.

However, as seen from the foundational perspective of Walley, the robust Bayesian approach has some serious defects, and sets of probability distributions are not an adequate foundation for a general theory of imprecise probability. There are many applications, of belief functions and possibility measures in particular, in which it is misleading to regard a set of probability distributions as a set of hypotheses about the "correct" probabilities, because it is meaningless to talk of "correct" probabilities.

6. On the relationship between Bayesian inference and evidence theory

Although many papers discuss the abstract relationships between different approaches to uncertainty management, such papers are less abundant that compare the precision and robustness of different approaches in specific application examples. Notable exceptions are some of the papers from the Sandia workshop, presented in [43] (see Section 7), as well as the paper [11] where Arnborg discusses the relationship between robust Bayesian inference and evidence theory, using Zadeh's ill-conditioned example [48] [49] to illustrate the effects on conclusions of different fusion rules, or rules of combination. Assuming evidence theory is being used, Arnborg notes that to obtain bodies of evidence, likelihoods and priors are needed, and therefore a robust Bayesian analysis based on these likelihoods and priors might be used to compare the conclusions emanating from the application of evidence theory with those from a purely Bayesian approach. The approach used in [11] assumes that impreciseness in conclusions is caused by impreciseness in sampling functions and priors.

The Dempster-Shafer (DS) combination rule [44] is computationally equivalent to allowing the operands as well as the result to be non-empty random sets. The combination of evidence - likelihood functions normalized so that they can be seen as probability distributions - and a prior over a finite space is done in [11] by component-wise multiplication followed by normalization. For precise beliefs, the resulting combination operation agrees with the DS and the MDS rules, the latter proposed by Fixsen and Mahler [50] and involving a re-weighting of the operands. The robust Bayesian version of this procedure would replace the probability distributions by sets of probability distributions, for example represented as DS belief functions.

Before continuing with the presentation of Arnborg's paper, it is appropriate to define the concept of *Choquet capacities of order 2* [46]:

Let O denote the set of possible outcomes under consideration. Suppose that lower probabilities $\underline{P}(A)$ are defined for all elements A in K, where K is a collection of subsets of O. Here, K is assumed to be an algebra. Lower probabilities determine conjugate upper probabilities through $\overline{P}(A) = 1 - \underline{P}(A^c)$, so it suffices to consider lower probabilities (upper and lower probabilities are also known as interval-valued or interval or non-additive probabilities). Let $\varnothing$ denote the empty set. Assume that $0 \le \underline{P}(A) \le 1$ for all A in K, $\underline{P}(\varnothing) = 0$ and $\underline{P}(O) = 1$. The lower probability $\underline{P}$ is said to be *2-monotone*, or a *Choquet capacity of order 2* or a *convex capacity*, when it also satisfies, whenever A and B are in K, $\underline{P}(A \vee B) + \underline{P}(A \wedge B) \ge \underline{P}(A) + \underline{P}(B)$.

Rounding. Arnborg shows how a set of distributions which is not a Choquet capacity of order 2 can be approximated by *rounding* it to a minimal Choquet capacity that contains it, and this rounded set can be represented by a DS-structure. This is a convex set of probability distributions which can be very compactly represented (typically by a few real numbers). Thus, imprecise distributions can, if constrained by rounding to Choquet capacities, be viewed as random sets. The random sets can be combined by taking the intersection of the participating random sets on condition that the result is non-empty (i.e., component-wise multiplication followed by normalization) and the resulting random set can be regarded as a Choquet capacity. Arnborg introduces the concepts of *robust Bayesian combination operator* and *rounded robust Bayesian combination operator* and notes that they are both monotone with respect to imprecision.

When interpreting DS-structures as Choquet capacities of order 2, it is highly desirable that the combination gives a capacity that is contained in the robust rule result. In fact, the MDS rule, viewed as a capacity, is contained in the robust Bayesian fusion result. This is not true in general for Dempster's rule, however. Furthermore, Arnborg notes that, unlike the robust and rounded robust Bayesian combination operators he proposes, the DS and MDS operators are not monotone with respect to imprecision. In fact, they either underestimate or eliminate imprecision, whereas the maximum entropy principle used in robust

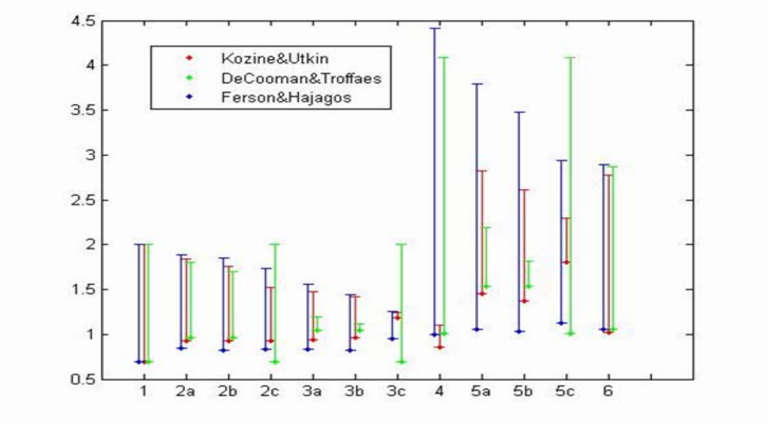

Figure 2. Results of three imprecise probability evaluation methods when applied to the six Sandia test problems.

Bayesianism can be given a rational game interpretation, and gives a different result in many cases.

Thus, evidence theory and robust Bayesianism are different in their conclusions. Arnborg concludes that "further work is required for understanding the basis for assessing uncertainty objectively, so that a given problem will not have incompatible solutions in the two frameworks. ... For higher level uncertainty management, dealing with quantities recognizable to users like military commanders, the need for clarity cannot be exaggerated.".

7 Imprecise probability methods and the Sandia challenge problems

The "Sandia challenge problems" were presented in [15] as addressing issues in the representation and aggregation of information concerning model parameters. The information can be of different types and may emanate from a number of sources, including measurements and expert opinion. The Sandia challenge problems embody several issues that concern all technologies for uncertainty propagation, whether probabilistic or not, distributional or not, approximate or rigorous. These issues are:

1. Aggregation of information from different sources (such as expert judgements).
2. Combination of probabilistic and non-probabilistic uncertainty.
3. Repetition of uncertain parameters.

The problem set is given by a simple algebraic system of the form $y = (a + b)^a$. Within this problem set, six main variants (and in some cases, subvariants, see [15]) were proposed:

Problem 1. a and b are contained in the closed real interval A and B, respectively.

Problem 2. a is contained in the closed interval A, and the information concerning b is given by n independent and equally credible sources. Each source specifies a closed interval B_i of possible values for b.

Problem 3. The information concerning both a and b is given by independent and equally credible sources of information, m sources for a and n sources for b.

Problem 4. a is contained in the closed interval A, and b is given by a log-normal distribution, $\ln b \sim N(\mu, \sigma)$. The value of the mean, μ, and the standard deviation, σ, are given by the closed intervals M and S, respectively.

Problem 5. The information concerning a is given by m independent and equally credible sources of information. Each source specifies a closed interval A_i that contains the value of a. The information concerning b is given by n independent and equally credible sources of information. Each source specifies closed intervals, M_j and S_j, of possible values of the mean and the standard deviation of the log-normal distribution.

Problem 6. a is contained in the closed interval A and b is given by a log-normal distribution. The values of both the mean and the standard deviation are precisely known.

Imprecise coherent probabilities. In [51], Kozine and Utkin use coherent imprecise probability models to estimate the results for the Sandia challenge problems. Partial knowledge is modeled by providing a set of admissible distributions that is smaller than all possible distributions. In all of the problems 1-6 expected values are sought. Since the source information is partial, i.e., the distributions are not precisely known, expected values can however only be found approximately as intervals. Such intervals are computed by formulating and solving analytically each problem as an optimization task.

Coherent lower previsions. In [47], de Cooman and Troffaes discuss why coherent lower previsions provide a good uncertainty model for solving generic uncertainty problems involving possibly conflicting expert information. They review the definition and meaning of important concepts in imprecise probability models, adding up to a concise and readable introduction to the subject. Finally, they apply their proposed optimization approach to the Sandia challenge problems, arguing that the theory of coherent lower previsions is eminently suited for solving the set of problems. Like Kozine and Utkin, their approach is by analytical optimization but the problems are formulated and solved differently.

Arithmetic with uncertain numbers. Ferson and Hajagos [52] solve the Sandia challenge problems using *probability bounds analysis*, an extension into probabilistic applications of the classical robust numerical methodology of *interval analysis* [53]. The inputs are expressed as interval bounds on cumulative distribution functions. Probability bounds analysis is based on work by Williamson and Downs [54], who developed an approach that computes rigorous bounds on the cumulative distribution functions of convolutions. According to [52], although probability bounds analysis does not prescribe a general solution for the question of how to aggregate information from disparate sources, it does offer a workable, albeit computationally intensive, strategy for handling repetitions of uncertain parameters in expressions. Using mathematical programming techniques, Berleant *et al*. independently derived and implemented arithmetical interval analysis algorithms to compute convolutions of bounded probability distributions, both with and without independence assumptions. The application of these methods to the Sandia challenge problems is reported in [55]. Results are of comparable quality to those in [52].

From the point of view of HLFB-DSS applications the Sandia problems are very simple. Nonetheless, the diagram in Figure 2 above shows limited quantitative agreement between the different approaches, indicating that also in this case, more comparative research is needed in order to eventually reach consensus on what imprecise probability methodologies are both feasible in practice and truly trustworthy.

8. Conclusion

From an engineering perspective of trustworthy information fusion, robust Bayesian concepts, and probably other kinds of imprecise probabilities as well, are likely to become of fundamental importance. Although robust Bayesian decision modeling is still an active area of research, it seems that methods and demonstrator systems have reached a level of maturity that is sufficient for inclusion in information fusion test applications. Robust sensor perception and target behavior submodels equipped with global reliability estimates to achieve robustness remain to be demonstrated, however. Few papers address these issues, and much more research is needed.

Emerging theories of imprecise probability indicate that a unified robust framework can be developed, in which several methodologies which were often considered incoherent and incompatible with Bayesian methods, such as belief or possibility measures, may be integrated. As illustrated by [11], such integration may require the development of methods of transforming and bounding of imprecise probability results in terms of robust Bayesian concepts. Judging from the sometimes widely differing attitudes of experts in these areas, however, a consentient unification of imprecise probability concepts in uncertainty management is probably not to be expected in the near future.

Noteworthy efforts are ongoing to include numerical approaches such as probability bounds analysis in safety-critical risk management applications.

In general it is very hard to solve all reliability issues before a new kind of system is fielded. When HLFB-DSS for a network-enabled defence are to be built, exploiting new sensor, communication, and modeling technology, this dilemma will become acute, and ways to plausibly demonstrate system reliability and robustness will have to be developed. In that undertaking, robust Bayesian analysis is likely to become the common yardstick.

References

[1] Llinas, J. (2001) Assessing the performance of multisensor fusion processes. In D. L. Hall and J. Llinas (eds.): *Handbook of Multisensor Data Fusion*, CRC Press, Boca Raton, LA, USA.

[2] Ahlberg, S., P. Hörling, K. Johansson, K. Jöred, H. Kjellström, C. Mårtenson, G. Neider, J. Schubert, P. Svenson, P. Svensson, and J. Walter (2006) An information fusion demonstrator for tactical intelligence processing in network-based defence. *Information Fusion*, in press.

[3] Looney, C. G. and L. R. Liang (2003) Cognitive situation and threat assessments of ground battlespaces. *Information Fusion* 4, 297-308.

[4] Sheridan, T. B. (1988) Trustworthiness of command and control systems. In *IFAC Man-Machine Systems*, 427-431.

[5] Bisantz, A. M., R. Finger, Y. Seong, and J. Llinas (1999) Human performance and data fusion based decision aids. In *Proc. 2nd Int. Conf. on Information Fusion*, 918-925.

[6] Pfautz, J., A. Fouse, T. Fichtl, A. Bisantz, E. Roth, and S. Madden (2005) The impact of meta-information on decision-making in intelligence operations. In *Proc. Human Factors and Ergonomics Society 49th Annual Meeting.*

[7] Falzon, L. (2006) Using Bayesian network analysis to support centre of gravity analysis in military planning, *European J. Operational Research*, 170(2) 629-643.

[8] Jensen, F. V. , T. D. Nielsen, and P. P. Shenoy (2006) Sequential influence diagrams: A unified asymmetry framework, *Int. J. Approx. Reasoning* 42(1-2) 101-118.

[9] Brynielsson, J. (2006) *A gaming perspective on command and control*, Doctoral thesis in computer science TRITA-CSC-A 2006:7, KTH, Stockholm.

[10] Johansson, L. R. M. and R. Suzic (2005) Particle-filter-based information acquisition for robust plan recognition, in *Proc. 8th Int. Conf. on Information Fusion*.

[11] Arnborg, S. (2004) Robust Bayesianism: Imprecise and paradoxical reasoning, in *Proc. 7th Int. Conf. on Information Fusion*, 407-414.

[12] Arnborg, S. Robust Bayesianism: Relation to evidence theory, *J. Advances in Information Fusion*, to appear.

[13] Forchini, G. and G. Hillier (2005) Ill-conditioned problems, Fisher information, and weak instruments, CEMMAP working paper CWP04/05, Univ. of Southampton, Southampton, UK.

[14] Demmel, J. W. (1987) On condition numbers and the distance to the nearest ill-posed problem, *Numerische Mathematik* 51, 251-289.

[15] Oberkampf, W. L., J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson (2004) Challenge problems: uncertainty in system response given uncertain parameters, in [43] 11-19.

[16] French, S. and D. Rios-Insua (2000) *Statistical Decision Theory*, Arnold, London.

[17] Berger, J. O. (1980) *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.

[18] Lindley, D. V. (1985) *Making decisions*. 2nd ed. Wiley Interscience, New York, NY, USA.

[19] Apostolakis, G.. (1990) The concept of probability in safety assessment of technological systems. *Science* 250, 1359-1364.

[20] Burdic, W. S. (1991) *Underwater acoustic system analysis*, Prentice Hall Inc., Upper Saddle River, NJ, USA.

[21] Biberman, L. M. (ed.) (2000) *Electro-optical imaging: system performance and modeling*, SPIE Press, Bellingham, WA, USA.

[22] Klein, L. A. (1997) *Millimeter-wave and infrared multisensor design and signal processing*, Artech House, Norwood, MA, USA.

[23] Horvitz, E. (1999) Principles of mixed-initiative user interfaces, in *Proc. ACM SIGCHI Conf. on Human Factors in Computing Systems (CHI '99)*, 159-166.

[24] Jiménez, A., S. Rios-Insua, and A. Mateos (2006) A generic multi-attribute analysis system, *Computers & Operations Research* 33(4), 1081-1101.

[25] Keefer, D. L., J. L. Corner, and C. W. Kirkwood (2002) Decision analysis applications in the operations research literature 1990-2001, Technical report, Arizona State University, USA.

[26] Ekenberg, L. and Torbiörnson, J. (2001) Second-order decision analysis, in *Int. J. Uncertainty, Fuzziness and Knowledge-based Systems* 9(1), 13-38.

[27] Larsson, A., J. Johansson, L. Ekenberg, and M. Danielsson (2005) Decision analysis with multiple objectives in a framework for evaluating imprecision, in *Int. J. Uncertainty, Fuzziness and Knowledge-based Systems* 13(5), 495-509.

[28] Berger, J. O. (1994) An overview of robust Bayesian analysis, TEST 3, 5-58.

[29] Walley, P. (1996) Measures of uncertainty in expert systems, *Artificial Intelligence* 83, 1-58.

[30] Kadane, J. B. and J. T. Chuang (1978) Stable decision problems, *Annals of Statistics* 4, 1095-1110.

[31] Ruggeri, F., D. Rios-Insua, and J. Martin (2004). Robust Bayesian analysis, in Dey (ed.) *Handbook of Statistics: Bayesian Analysis*, North-Holland, Amsterdam.

[32] Berger, J. O., D. Rios-Insua, F. Ruggeri (2000) Bayesian robustness, in [33], 1-32.

[33] Rios-Insua, D. and F. Ruggeri (eds.) (2000) *Robust Bayesian Analysis*, Lecture notes in statistics 152, Springer-Verlag, New York.

[34] Saltelli, A., K. Chan, and E. M. Scott, Eds (2000) *Sensitivity Analysis*, Wiley, West Sussex, UK.

[35] Kadane, J., G. Salinetti, and C. Srinivasan (2000) Stability of Bayes decisions and applications, in [33], 187-196.

[36] Kleijnen, J. P. C. and R. G. Sargent (2000) A methodology for fitting and validating metamodels in simulation, *European J. Operational Research* 120(1), 14-29.

[37] Frey, H. C. and S. R. Patil (2002) Identification and review of sensitivity analysis methods, *Risk Analysis* 22(3), 553-578.

[38] Kleijnen, J. P. C. (1998) Experiment design for sensitivity analysis, optimization, and validation of simulation models. In J. Banks (ed), *The Handbook of Simulation*, Wiley, New York.

[39] Balci, O. (2001) A methodology for certification of modeling and simulation applications, *ACM Trans. on Modeling and Computer-based Simulation* 11(4), 352-377.

[40] Johansson, L. R. M (2006) *Large-scale information acquisition for data and information fusion*, Doctoral thesis in computer science TRITA-CSC-A 2006:2, KTH, Stockholm.

[41] Cozman, F. C. and Moral, S. (eds.) (2000) Reasoning with imprecise probabilities, special issue of *Int. J. Approx. Reasoning* 24, 121-299.

[42] Zaffalon, M. (ed.) (2005) Credibility via imprecise probability, special issue of *Int. J. Approx. Reasoning* 39(2-3), 119-324.

[43] Helton, J. C. and W. L. Oberkampf (eds.) (2004) Alternative representations of epistemic uncertainty, *Reliability Engineering & System Safety* 85, 1-369.

[44] Shafer, G. (1976) *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, USA.

[45] Zadeh, L. A. (1978) Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* 1, 3-28.

[46] Walley, P. (2000) Towards a unified theory of imprecise probability, in [41], 125-148.

[47] de Cooman, G. and M. C. M. Troffaes (2004) Coherent lower previsions in systems modelling: products and aggregation rules, in [43], 113-134.

[48] Zadeh, L. A. (1986) Is probability theory sufficient for dealing with uncertainty in AI: a negative view, in: L. N. Kanal, J. F. Lemmer (eds.) *Uncertainty in Artificial Intelligence*, North-Holland, Amsterdam, 103-116.

[49] Haenni, R. (2005) Shedding new light on Zadeh's criticism of Dempster's rule of combination, *Proc. of the 8th Int. Conf. on Information Fusion*.

[50] Fixsen, D. and R. P. S. Mahler (1997) The modified Dempster-Shafer approach to classification. *IEEE Trans. SMC-A*, 27(1), 96-104.

[51] Kozine, I. O. and L. V. Utkin (2004) An approach to combining unreliable pieces of evidence and their propagation in a system response analysis, in [43] 103-112.

[52] Ferson, S. and J. G. Hajagos (2004) Arithmetic with uncertain numbers: rigorous and (often) best possible answers, in [43] 135-152.

[53] Moore, R. E. (1979) *Methods and applications of interval analysis*, SIAM, Philadelphia, PA, USA.

[54] Williamson, R. C. and T. Downs (1990) Probabilistic arithmetic I: numerical methods for calculating convolutions and dependency bounds, *Int. J. Approx. Reasoning* 4, 89-158.

[55] Berleant, D. and J. Zhang (2004) Representation and problem solving with Distribution Envelope Determination (DEnv), in [43] 153-168.